

Analisis Sentimen Konsumen pada Rumah Makan di Mataram Menggunakan Algoritma *K-Nearest Neighbor*, *Naïve Bayes*, dan *Support Vector Machine*

Ameylan Verina Tabun^{1*}, Hairani² Dadang Priyanto³

^{1*,2,3} Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Bumigora, Kota Mataram, Nusa Tenggara Barat, Indonesia.

Corresponding Email: ameylanel@gmail.com^{1*}

Histori Artikel:

Dikirim 12 Agustus 2026; *Diterima dalam bentuk revisi* 27 Agustus 2026; *Diterima* 06 September 2026; *Diterbitkan* 28 September 2026. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

Abstrak

Ulasan konsumen pada platform digital dapat digunakan untuk mengetahui persepsi pelanggan terhadap kualitas makanan, pelayanan, harga, dan kenyamanan rumah makan. Penelitian ini menggunakan 1.006 ulasan dari 18 rumah makan di Kota Mataram yang diperoleh melalui *web scraping* pada TripAdvisor. Setelah pembersihan dan seleksi data, 780 ulasan digunakan sebagai dataset akhir. Penelitian bertujuan menganalisis sentimen ulasan konsumen serta membandingkan kinerja algoritma *K-Nearest Neighbor* (K-NN), *Naïve Bayes* (NB), dan *Support Vector Machine* (SVM) dalam klasifikasi sentimen positif, negatif, dan netral. Metode yang digunakan adalah kuantitatif dengan eksperimen komparatif. Tahapan penelitian mencakup pengumpulan data, *preprocessing*, pelabelan sentimen berbasis *lexicon* dan rating, ekstraksi fitur menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF), klasifikasi, dan evaluasi dengan *confusion matrix*. Pengujian dilakukan pada pembagian data 70:30, 80:20, dan 90:10 dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil menunjukkan bahwa SVM menghasilkan akurasi tertinggi pada seluruh skenario pengujian. Performa terbaik diperoleh pada pembagian data 80:20 dengan *accuracy* 87,1%, *precision* 87,0%, *recall* 87,0%, dan *F1-score* 86,0%.

Kata Kunci: Analisis Sentimen; TripAdvisor; *Text Mining*; TF-IDF; K-NN; *Naïve Bayes*; SVM.

Abstract

Consumer reviews on digital platforms can be used to identify customer perceptions of food quality, service, price, and restaurant comfort. This study used 1,006 reviews from 18 restaurants in Mataram City collected through web scraping from TripAdvisor. After data cleaning and selection, 780 reviews were used as the final dataset. The study aimed to analyze consumer review sentiment and compare the performance of K-Nearest Neighbor (K-NN), Naïve Bayes (NB), and Support Vector Machine (SVM) algorithms in classifying positive, negative, and neutral sentiments. A quantitative approach with a comparative experimental method was employed. The research stages included data collection, preprocessing, sentiment labeling based on lexicon and rating, feature extraction using Term Frequency-Inverse Document Frequency (TF-IDF), classification, and evaluation using a confusion matrix. Tests were conducted using 70:30, 80:20, and 90:10 data splits with accuracy, precision, recall, and F1-score as evaluation metrics. The results showed that SVM achieved the highest accuracy across all testing scenarios. The best performance was obtained with the 80:20 split, reaching 87.1% accuracy, 87.0% precision, 87.0% recall, and 86.0% F1-score.

Keyword: Sentiment Analysis; TripAdvisor; *Text Mining*; TF-IDF; K-NN; *Naïve Bayes*; SVM.

1. Pendahuluan

Perkembangan teknologi digital telah mengubah cara konsumen mencari, membandingkan, dan mengevaluasi produk maupun layanan. Saat ini, ulasan daring menjadi salah satu sumber informasi penting karena memungkinkan konsumen menyampaikan pengalaman secara langsung setelah menggunakan suatu layanan. Pada sektor kuliner, ulasan memuat penilaian terhadap berbagai aspek, seperti rasa makanan, harga, pelayanan, kebersihan, suasana tempat, dan kepuasan secara keseluruhan. Informasi dalam ulasan memiliki nilai yang tinggi karena mencerminkan pandangan pengguna secara nyata terhadap kualitas rumah makan. Platform seperti TripAdvisor banyak dimanfaatkan konsumen untuk memberikan penilaian, sehingga menghasilkan data dalam jumlah besar yang dapat digunakan dalam penelitian berbasis teks. Penelitian sebelumnya juga menunjukkan bahwa ulasan pelanggan pada layanan dan produk dapat dimanfaatkan untuk memahami tingkat kepuasan serta kecenderungan opini pengguna (Firdaus *et al.*, 2024; Pratama & Fajri, 2025).

Data ulasan pada platform digital umumnya disajikan dalam bentuk teks bebas. Bentuk data tersebut bersifat tidak terstruktur dan memuat variasi bahasa yang luas. Konsumen dapat menulis ulasan dengan gaya formal, informal, singkat, panjang, menggunakan singkatan, bahkan mencampur bahasa Indonesia dan bahasa Inggris. Kondisi tersebut membuat identifikasi opini secara manual menjadi kurang efisien, terutama ketika jumlah ulasan terus bertambah. Selain memerlukan waktu yang lama, pembacaan manual juga berpotensi menimbulkan perbedaan penilaian antarpeneliti. Oleh karena itu, diperlukan pendekatan komputasional yang mampu mengolah teks secara otomatis agar informasi di dalamnya dapat dianalisis secara sistematis. Pendekatan semacam itu telah banyak digunakan pada kajian ulasan pelanggan di berbagai bidang, termasuk restoran, layanan digital, dan platform media sosial (Ichsansyah & Subektiningsih, 2026; Damaryana, 2025).

Analisis sentimen merupakan salah satu pendekatan yang banyak digunakan untuk mengidentifikasi kecenderungan opini dari teks. Metode tersebut bekerja dengan mengelompokkan ulasan ke dalam kategori positif, negatif, atau netral. Dalam bidang kuliner, analisis sentimen dapat membantu mengetahui cara konsumen menilai rumah makan berdasarkan pengalaman mereka. Hasil analisis dapat digunakan oleh pemilik usaha untuk memahami aspek yang perlu dipertahankan maupun diperbaiki. Bagi peneliti, hasil klasifikasi sentimen dapat menjadi dasar dalam menilai pola opini konsumen pada layanan restoran. Penelitian terdahulu menunjukkan bahwa analisis sentimen mampu memberikan gambaran yang cukup baik terhadap persepsi pelanggan pada layanan makanan dan restoran, termasuk pada aspek *food*, *service*, dan *value* (Singgalen, 2023; Rukmana *et al.*, 2025). Dengan begitu, analisis sentimen tidak hanya berguna dalam pengolahan data teks, tetapi juga memiliki nilai praktis dalam mendukung pengambilan keputusan.

Sejumlah penelitian terdahulu menunjukkan bahwa analisis sentimen dapat diterapkan pada ulasan restoran dengan hasil yang beragam. Salsabillah *et al.* (2024) membandingkan *Support Vector Machine* dan *Naïve Bayes* dalam klasifikasi sentimen ulasan rumah makan. Sifa Amalia *et al.* (2021) membandingkan *Support Vector Machine* dan *K-Nearest Neighbor* pada data ulasan pelanggan restoran. Prasetyo *et al.* (2023) juga menerapkan *Naïve Bayes* dan *K-Nearest Neighbor* untuk analisis serta kategorisasi ulasan pelanggan pada cafe. Hasil penelitian tersebut menunjukkan bahwa teks ulasan konsumen dapat diolah menjadi informasi melalui tahapan analisis sentimen. Selain itu, pemilihan algoritma memengaruhi performa klasifikasi yang dihasilkan. Setiap metode memiliki keunggulan dan keterbatasan yang bergantung pada karakteristik data (Purnawibawa *et al.*, 2022; Pratama & Fajri, 2025).

Kajian serupa juga dilakukan pada berbagai objek di luar sektor kuliner. Syahril Dwi Prasetyo *et al.* (2023) menerapkan *Naïve Bayes* dan KNN untuk analisis sentimen relokasi ibukota Nusantara. Muliawan dan Dazki (2023) membandingkan tiga algoritma, yaitu *Naïve Bayes*, KNN, dan *Random Forest*, untuk analisis sentimen isu relokasi ibu kota. Nabila *et al.* (2025) mengkaji ulasan Starbucks dengan menerapkan KNN dan SVM. Selain itu, Damaryana (2025) menganalisis komentar YouTube mengenai program Makan Bergizi Gratis menggunakan SVM dan KNN. Hasil dari

berbagai penelitian tersebut memperlihatkan bahwa metode klasifikasi teks dapat digunakan secara luas pada beragam jenis data ulasan maupun opini publik.

Walaupun penelitian mengenai analisis sentimen pada ulasan restoran telah banyak dilakukan, kajian pada rumah makan lokal di Kota Mataram masih belum banyak ditemukan, khususnya yang memanfaatkan data dari TripAdvisor. Padahal, rumah makan di daerah tersebut memiliki variasi layanan dan karakteristik pelanggan yang dapat menghasilkan pola ulasan berbeda. Faktor lokasi, jenis makanan, harga, pelayanan, dan ekspektasi konsumen dapat memengaruhi isi ulasan yang ditulis. Selain itu, penggunaan bahasa pada ulasan lokal juga dapat memiliki ciri khas tertentu yang berbeda dari wilayah lain. Perbedaan karakteristik tersebut menunjukkan bahwa hasil penelitian dari daerah lain belum tentu dapat diterapkan secara langsung pada data rumah makan di Kota Mataram. Karena itu, diperlukan kajian khusus yang berfokus pada ulasan konsumen di wilayah tersebut.

Penelitian ini menggunakan tiga algoritma klasifikasi, yaitu *K-Nearest Neighbor* (K-NN), *Naïve Bayes*, dan *Support Vector Machine* (SVM). Ketiga algoritma dipilih karena memiliki pendekatan yang berbeda dalam klasifikasi teks. K-NN bekerja dengan mengukur kedekatan data uji terhadap data latih terdekat dan menentukan kelas berdasarkan mayoritas tetangga. Pendekatan tersebut sederhana dan mudah dipahami, namun hasilnya sangat bergantung pada jumlah tetangga dan distribusi data. *Naïve Bayes* menggunakan prinsip probabilitas untuk menghitung peluang suatu teks masuk ke dalam kelas tertentu berdasarkan kemunculan fitur. Metode tersebut sering digunakan pada klasifikasi teks karena efisien dan mampu bekerja baik pada data berukuran besar. Sementara itu, SVM berupaya menemukan batas pemisah terbaik antar kelas dengan membentuk *hyperplane* optimal. Metode tersebut dikenal memiliki kinerja yang baik pada data berdimensi tinggi seperti teks hasil ekstraksi fitur. Sejumlah studi juga menunjukkan bahwa perbandingan KNN dan SVM dapat menghasilkan performa yang berbeda bergantung pada data yang digunakan (Purnawibawa *et al.*, 2022; Nabila *et al.*, 2025).

Perbandingan ketiga algoritma tersebut penting dilakukan karena karakteristik data ulasan bersifat dinamis dan memiliki variasi kata yang tinggi. Setiap algoritma dapat menunjukkan performa berbeda ketika diterapkan pada dataset yang sama. Oleh sebab itu, pengujian terhadap K-NN, *Naïve Bayes*, dan SVM diperlukan untuk mengetahui metode yang paling sesuai dalam mengklasifikasikan sentimen ulasan rumah makan di Kota Mataram. Evaluasi dilakukan berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score* agar hasil yang diperoleh tidak hanya menunjukkan tingkat ketepatan umum, tetapi juga keseimbangan antara kemampuan model dalam mengenali kelas tertentu dan keseluruhan prediksi yang dihasilkan.

Penelitian ini diharapkan dapat memberikan gambaran mengenai kecenderungan sentimen konsumen terhadap rumah makan di Kota Mataram berdasarkan ulasan pada TripAdvisor. Selain itu, hasil pengujian dapat menjadi referensi dalam pemilihan metode klasifikasi teks yang efektif pada data ulasan berbahasa Indonesia. Dari sisi akademik, penelitian ini dapat memperkaya kajian mengenai penerapan *text mining* pada bidang kuliner. Dari sisi praktis, hasil penelitian dapat membantu pelaku usaha memahami penilaian konsumen terhadap layanan yang diberikan sehingga dapat menjadi bahan pertimbangan dalam peningkatan kualitas usaha.

2. Metode Penelitian

2.1. Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen komparatif. Tahapan penelitian terdiri atas pengumpulan data, *preprocessing*, seleksi data, pelabelan sentimen, ekstraksi fitur menggunakan TF-IDF, pembagian data, proses klasifikasi, dan evaluasi model. Pendekatan komparatif semacam ini umum digunakan dalam penelitian analisis sentimen untuk membandingkan performa beberapa algoritma pada dataset yang sama.

2.2 Pengumpulan dan Seleksi Data

Data penelitian diperoleh melalui proses *web scraping* pada platform TripAdvisor pada periode 7 November 2025 hingga 7 Juni 2026. Dataset awal terdiri atas 1.006 ulasan yang berasal dari 18 rumah makan di Kota Mataram. Pada tahap pemeriksaan awal, ditemukan 24 data kosong sehingga data tersebut tidak dapat digunakan. Setelah penghapusan data kosong, jumlah data menjadi 982 ulasan. Selanjutnya, dilakukan seleksi kelayakan data sebelum proses pelabelan dan pemodelan. Berdasarkan proses seleksi, dataset akhir yang digunakan dalam eksperimen berjumlah 780 ulasan. Dataset tersebut digunakan sebagai dasar proses pelabelan, ekstraksi fitur, pelatihan, dan pengujian model. Tahap seleksi data diperlukan untuk memastikan kualitas data teks yang akan dianalisis, terutama karena ulasan pengguna sering kali mengandung teks yang tidak relevan atau tidak lengkap (Samudra *et al.*, 2025; Rahayu *et al.*, 2022).

Tabel 1. Jumlah Data Hasil Seleksi

Tahap	Jumlah Data
Data hasil <i>web scraping</i>	1.006
Data kosong yang dihapus	24
Data setelah <i>cleaning</i> awal	982
Dataset akhir yang digunakan dalam eksperimen	780

2.3. Preprocessing Data

Tahap *preprocessing* dilakukan untuk mengubah teks ulasan menjadi data yang lebih bersih dan terstruktur. Tahapan *preprocessing* terdiri atas *cleaning*, *case folding*, *normalization*, *tokenizing*, *stopword removal*, dan *stemming*. *Cleaning* digunakan untuk menghapus elemen yang tidak relevan. *Case folding* menyeragamkan huruf menjadi huruf kecil. *Normalization* memperbaiki kata tidak baku atau singkatan. *Tokenizing* memecah teks menjadi token. *Stopword removal* menghapus kata umum yang kurang informatif. *Stemming* mengubah kata menjadi bentuk dasar, misalnya kata *makanan* menjadi *makan* dan *pelayanan* menjadi *layan*. Tahapan tersebut lazim diterapkan pada penelitian analisis sentimen berbasis ulasan untuk meningkatkan kualitas data masukan sebelum proses klasifikasi (Sari *et al.*, 2023; Samudra *et al.*, 2025).

2.4. Pelabelan Sentimen

Pelabelan sentimen dilakukan menggunakan pendekatan berbasis *lexicon* yang dikombinasikan dengan informasi rating TripAdvisor. Rating 4 dan 5 dikategorikan sebagai sentimen positif, rating 3 sebagai netral, sedangkan rating 1 dan 2 sebagai negatif. Penggunaan *lexicon* dilakukan untuk mengidentifikasi kecenderungan kata yang memiliki makna positif dan negatif dalam teks ulasan. Pendekatan pelabelan berdasarkan ulasan pengguna dan skor penilaian juga digunakan dalam berbagai penelitian sentimen sebelumnya karena mampu memberikan dasar klasifikasi yang lebih terarah (Rahat *et al.*, 2020; Samudra *et al.*, 2025).

Tabel 2. Rekapitulasi Seleksi Dataset Penelitian

Sentimen	Jumlah Data	Persentase
Positif	508	65,1%
Negatif	150	19,2%
Netral	122	15,6%
Total	780	100%

2.5. Ekstraksi Fitur Menggunakan TF-IDF

Teks ulasan diubah menjadi representasi numerik menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). TF-IDF memberikan bobot pada setiap kata berdasarkan frekuensi kemunculannya dalam suatu dokumen dan tingkat kelangkaan kata tersebut dalam keseluruhan dataset. Hasil pembobotan kemudian digunakan sebagai fitur masukan bagi algoritma klasifikasi.

Teknik TF-IDF banyak digunakan pada analisis sentimen karena mampu merepresentasikan teks secara efektif pada data ulasan yang bersifat tinggi dimensi (Sari *et al.*, 2023; Rahayu *et al.*, 2022).

2.6. *K-Nearest Neighbor*

K-Nearest Neighbor digunakan untuk mengklasifikasikan data berdasarkan kedekatan dengan sejumlah data latih. Pada data berbentuk TF-IDF, klasifikasi dilakukan berdasarkan kemiripan antarrepresentasi vektor dokumen. Nilai parameter k perlu disesuaikan dengan konfigurasi implementasi yang digunakan pada eksperimen agar dapat direplikasi secara tepat. KNN telah banyak digunakan pada analisis sentimen ulasan karena sederhana dan cukup efektif untuk data teks dengan representasi numeric.

2.7. *Naïve Bayes*

Model *Naïve Bayes* yang digunakan adalah *Multinomial Naïve Bayes* karena sesuai untuk data berbentuk fitur non-negatif hasil representasi teks. Model menghitung probabilitas setiap kelas berdasarkan fitur yang terdapat pada dokumen dan menentukan kelas dengan probabilitas terbesar sebagai hasil prediksi. Dalam beberapa penelitian, *Naïve Bayes* menunjukkan kinerja yang kompetitif pada klasifikasi sentimen berbasis ulasan, terutama karena proses komputasinya yang sederhana dan efisien (Rahat *et al.*, 2020).

2.8. *Support Vector Machine*

Support Vector Machine digunakan untuk mencari *hyperplane* optimal yang dapat memisahkan data antar kelas. Penelitian ini menggunakan SVM dengan *Linear Kernel* dan pendekatan *One-vs-Rest* untuk menangani tiga kelas sentimen. *Linear Kernel* sesuai dengan karakteristik data teks hasil TF-IDF yang memiliki dimensi tinggi dan bersifat *sparse*. SVM juga banyak digunakan dalam analisis sentimen ulasan karena mampu memberikan hasil yang baik pada berbagai jenis dataset teks.

2.9. Pembagian Data dan Evaluasi

Dataset dibagi menggunakan tiga skenario, yaitu 70:30, 80:20, dan 90:10. Evaluasi dilakukan menggunakan *confusion matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. *Accuracy* mengukur proporsi keseluruhan prediksi yang benar, *precision* mengukur ketepatan prediksi, *recall* mengukur kemampuan model menemukan data yang benar dari suatu kelas, dan *F1-score* merupakan rata-rata harmonis antara *precision* dan *recall*. Evaluasi dengan metrik tersebut umum digunakan dalam penelitian komparatif analisis sentimen untuk melihat performa antaralgoritma secara lebih menyeluruh.

3. Hasil dan Pembahasan

3.1 Hasil

3.1.1 Hasil *Preprocessing* dan Pelabelan

Tahap *preprocessing* menghasilkan teks ulasan yang lebih terstruktur dan siap digunakan dalam proses pembobotan dan klasifikasi. Dataset akhir yang digunakan berjumlah 780 ulasan, terdiri atas 508 ulasan positif, 150 ulasan netral, dan 122 ulasan negatif. Dominasi kelas positif menunjukkan bahwa sebagian besar ulasan memiliki kecenderungan positif. Kondisi ini perlu diperhatikan karena ketidakseimbangan distribusi kelas dapat memengaruhi kemampuan model dalam mengenali kelas minoritas.

3.1.2 Hasil TF-IDF

Hasil pembobotan TF-IDF menunjukkan bahwa kata *makan* memiliki nilai tertinggi, diikuti oleh *kurang*, *layan*, *baik*, dan *cukup*. Urutan tersebut memperlihatkan bahwa ulasan konsumen paling sering berputar pada makanan dan pelayanan. Kata *kurang* yang cukup menonjol juga

mengisyaratkan adanya penilaian kritis dari pengguna. Temuan ini membantu model mengenali pola bahasa yang paling sering muncul dalam ulasan.

Tabel 3. Hasil Pembobotan TF-IDF

No	Kata	Nilai TF-IDF
1	makan	81
2	kurang	77
3	layan	63
4	baik	49
5	cukup	48
6	enak	46
7	tempat	43
8	alam	41
9	rasa	40
10	nyaman	38

Nilai pada tabel menunjukkan tingkat kepentingan relatif kata dalam keseluruhan dataset berdasarkan hasil pembobotan TF-IDF. Nilai *makan* sebesar 81 tidak menunjukkan bahwa kata tersebut muncul sebanyak 81 kali, melainkan merupakan nilai pembobotan atau agregasi yang menunjukkan kontribusi kata terhadap representasi fitur. Kata *makan* menggambarkan konteks utama objek penelitian, kata *kurang* dapat menunjukkan adanya evaluasi terhadap kekurangan, sedangkan kata *layan* menunjukkan pentingnya aspek pelayanan dalam ulasan konsumen.

3.1.3 Hasil Evaluasi Model

Hasil evaluasi model menunjukkan bahwa SVM memperoleh accuracy tertinggi pada seluruh skenario pengujian, yakni 85% pada pembagian 70:30, 87% pada 80:20, dan 85% pada 90:10. K-NN dan Naïve Bayes masih berada di bawah SVM pada tiap skenario. Performa terbaik SVM terdapat pada pembagian data 80:20 dengan accuracy 87,1%, precision 87,0%, recall 87,0%, dan F1-score 86,0%. Hasil ini menunjukkan bahwa SVM lebih stabil dalam mengklasifikasikan sentimen ulasan.

Tabel 4. Hasil Evaluasi Model

Pembagian Data	K-NN	Naïve Bayes	SVM
70:30	76%	75%	85%
80:20	76%	77%	87%
90:10	79%	79%	85%

SVM memperoleh *accuracy* tertinggi pada seluruh skenario pengujian. Performa terbaik diperoleh pada *split* 80:20 dengan *accuracy* sebesar 87,1%, *precision* 87,0%, *recall* 87,0%, dan *F1-score* 86,0%.

3.1.4 Perbandingan Performa Algoritma

SVM memperoleh performa terbaik pada seluruh skenario pengujian. Salah satu faktor yang dapat menjelaskan hasil tersebut adalah karakteristik data TF-IDF yang berdimensi tinggi dan *sparse*. SVM dengan *Linear Kernel* dapat bekerja efektif dalam kondisi tersebut. K-NN bergantung pada pengukuran kedekatan antardokumen, sedangkan *Naïve Bayes* menggunakan pendekatan probabilistik dengan asumsi independensi fitur. Hasil tersebut menunjukkan bahwa SVM lebih sesuai dengan dataset dan representasi fitur yang digunakan dalam penelitian. Namun, hasil tersebut tidak berarti bahwa SVM selalu menjadi model terbaik pada semua dataset analisis sentimen karena performa model dapat dipengaruhi oleh jumlah data, distribusi kelas, *preprocessing*, pelabelan, parameter, dan representasi fitur.

3.2 Pembahasan

SVM memberikan performa paling baik dibandingkan K-NN dan *Naïve Bayes* pada klasifikasi sentimen ulasan rumah makan di Kota Mataram. Temuan ini sejalan dengan penelitian Sifa Amalia

et al. (2021) yang juga menemukan bahwa SVM mampu memberikan hasil yang lebih baik dalam analisis sentimen review pelanggan restoran dibandingkan K-NN. Hasil serupa juga ditunjukkan oleh Pratama dan Fajri (2025), yang menyatakan bahwa SVM lebih unggul daripada KNN dalam prediksi kepuasan pelanggan. Pada data teks yang berdimensi tinggi seperti TF-IDF, SVM cenderung lebih stabil karena mampu membentuk pemisah kelas yang lebih tegas.

Dominasi sentimen positif pada dataset penelitian ini memperlihatkan bahwa mayoritas konsumen memberikan penilaian yang baik terhadap rumah makan yang diteliti. Pola ini juga ditemukan pada kajian Singgalen (2023), yang menelaah aspek *food, services, and value* pada restoran dan rumah makan di Makassar berdasarkan ulasan Tripadvisor. Hasil tersebut menunjukkan bahwa ulasan konsumen sering kali berkisar pada pengalaman langsung terhadap rasa makanan, pelayanan, dan kenyamanan tempat. Dalam penelitian ini, kata *makan, layan, dan kurang* menjadi fitur dengan bobot tinggi, sehingga memperkuat dugaan bahwa tiga aspek tersebut paling sering memengaruhi isi ulasan.

Jika dibandingkan dengan penelitian Prasetyo *et al.* (2023), yang menggunakan *Naïve Bayes* dan K-NN pada review pelanggan cafe, hasil penelitian ini menunjukkan bahwa karakteristik data sangat memengaruhi performa algoritma. Pada beberapa dataset, *Naïve Bayes* dan K-NN dapat menghasilkan akurasi yang cukup baik, tetapi pada data ulasan rumah makan di Kota Mataram, SVM tetap lebih unggul. Hal ini sejalan dengan Purnawibawa *et al.* (2022), yang menunjukkan bahwa SVM sering lebih efektif dibandingkan KNN pada data berfitur numerik hasil ekstraksi teks, karena KNN lebih sensitif terhadap jarak antar data dan distribusi kelas.

Hasil pada skenario pembagian data 80:20 yang menghasilkan performa terbaik juga menunjukkan bahwa keseimbangan antara data latih dan data uji berpengaruh terhadap kualitas model. Temuan ini konsisten dengan penelitian Syahril Dwi Prasetyo *et al.* (2023) dan Muliawan dan Dazki (2023), yang memperlihatkan bahwa perbedaan algoritma dan pembagian data dapat menghasilkan tingkat akurasi yang berbeda pada analisis sentimen. Dengan kata lain, tidak ada satu model yang selalu unggul di semua kondisi, karena hasil sangat dipengaruhi oleh ukuran data, karakteristik fitur, dan distribusi kelas.

Dari sisi manfaat, hasil penelitian ini memperlihatkan bahwa analisis sentimen dapat digunakan sebagai alat bantu untuk membaca penilaian konsumen secara lebih terarah. Firdaus *et al.* (2024) menjelaskan bahwa analisis sentimen memiliki nilai praktis dalam mendukung pengambilan keputusan, termasuk pada bidang layanan dan kebijakan. Dalam penelitian ini, informasi dari ulasan dapat membantu pemilik rumah makan mengenali aspek yang perlu dipertahankan maupun diperbaiki, terutama pada makanan, pelayanan, dan kenyamanan. Namun, hasil yang diperoleh tetap perlu dibaca dengan hati-hati karena dominasi kelas positif dan keterbatasan dataset dapat memengaruhi kemampuan model dalam mengenali sentimen minoritas.

Penelitian Damaryana (2025) pada komentar YouTube tentang program Makan Bergizi Gratis juga menunjukkan bahwa KNN dan SVM dapat digunakan pada data opini publik, tetapi hasil terbaik tetap bergantung pada karakteristik data yang dianalisis. Hal ini memperkuat bahwa pemilihan algoritma tidak bisa dilepaskan dari bentuk data dan tujuan analisis. Oleh sebab itu, pada penelitian ini, SVM dapat dinilai paling sesuai untuk dataset ulasan rumah makan di Kota Mataram, meskipun hasil tersebut tetap terbuka untuk diuji ulang pada dataset yang lebih besar dan beragam.

4. Kesimpulan dan Saran

Penelitian ini melakukan analisis sentimen terhadap ulasan konsumen rumah makan di Kota Mataram menggunakan K-Nearest Neighbor, *Naïve Bayes*, dan *Support Vector Machine*. Data awal berjumlah 1.006 ulasan dari 18 rumah makan. Setelah proses pembersihan dan seleksi data, sebanyak 780 ulasan digunakan sebagai dataset akhir dalam proses pelabelan, ekstraksi fitur, pelatihan, dan pengujian model. Hasil pelabelan menunjukkan bahwa sentimen positif merupakan kelas dominan dengan 508 ulasan, diikuti sentimen netral sebanyak 150 ulasan dan sentimen negatif sebanyak 122 ulasan. Hasil TF-IDF menunjukkan bahwa kata *makan, kurang, dan layan* merupakan

beberapa fitur dengan nilai pembobotan tertinggi. Hasil perbandingan menunjukkan bahwa SVM memperoleh *accuracy* tertinggi pada seluruh skenario pengujian. Performa terbaik diperoleh pada pembagian data 80:20 dengan *accuracy* 87,1%, *precision* 87,0%, *recall* 87,0%, dan *F1-score* 86,0%. Berdasarkan dataset dan skenario pengujian yang digunakan dalam penelitian ini, SVM merupakan algoritma dengan performa terbaik untuk klasifikasi sentimen ulasan konsumen rumah makan di Kota Mataram. Hasil penelitian ini dapat dimanfaatkan sebagai informasi awal untuk membantu pemilik rumah makan memahami persepsi konsumen terhadap makanan, pelayanan, dan kenyamanan. Penelitian selanjutnya dapat menggunakan dataset yang lebih besar dan beragam, proses pelabelan manual, strategi penanganan ketidakseimbangan kelas, serta membandingkan algoritma *machine learning* dan pendekatan *deep learning* lainnya.

5. Daftar Pustaka

- Damaryana, L. (2025). *Analisis sentimen komentar YouTube terhadap program Makan Bergizi Gratis menggunakan algoritma Support Vector Machine dan K-Nearest Neighbor* [Doctoral dissertation, Universitas Nasional].
- Firdaus, A. A., Saputro, J. S., Anwar, M., Adriyanto, F., Maghfiroh, H., Ma'arif, A., ... Hidayat, R. (2024). Application of sentiment analysis as an innovative approach to policy making: A review. *Journal of Robotics and Control (JRC)*, 5(6), 1784–1798. <https://doi.org/10.18196/jrc.v5i6.22573>
- Ichsansyah, R. F. P., & Subektiningsih, S. (2026). Perbandingan algoritma Naïve Bayes dan Random Forest dalam klasifikasi sentimen ulasan pengguna Kredivo di Play Store. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 13(2), 297–308.
- Muliawan, J., & Dazki, E. (2023). Sentiment analysis of Indonesia's capital city relocation using three algorithms: Naïve Bayes, KNN, and Random Forest. *Jurnal Teknik Informatika (JUTIF)*, 4(5), 1227–1236. <https://doi.org/10.52436/1.jutif.2023.4.5.1436>
- Nabila, J., Putra, I. A. D. B., & Wijayanto, H. (2025). Sentiment analysis on Starbucks reviews: Implementation of K-Nearest Neighbors and Support Vector Machine. *Infact: International Journal of Computers*, 9(02), 79–86. <https://doi.org/10.61179/infact.v9i02.761>
- Prasetyo, S. D., Hilabi, S. S., & Nurapriani, F. (2023). Analisis sentimen relokasi ibukota Nusantara menggunakan algoritma Naïve Bayes dan KNN. *Jurnal KomtekInfo*, 10(1), 1–7. <https://doi.org/10.35134/komtekinfo.v10i1.330>
- Pratama, S. R., & Fajri, I. N. (2025). Comparison of K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) algorithms in predicting customer satisfaction. *Journal of Computer Science and Informatics Engineering*, 4(3), 135–146. <https://doi.org/10.55537/cosie.v4i3.1160>
- Purnawibawa, I. W. A., Purnama, I. N., & Wijaya, I. N. Y. A. (2022). Komparasi algoritme K-Nearest Neighbors dan Support Vector Machines dalam prediksi layanan produk ICONNET. *Progresif: Jurnal Ilmiah Komputer*, 18(2), 271–282. <https://doi.org/10.35889/progresif.v18i2.894>
- Rahat, A. M., Kahir, A., & Masum, A. K. M. (2020). Comparison of Naive Bayes and SVM algorithm based on sentiment analysis using review dataset. In *Proceedings of the 2019 8th International Conference on System Modeling and Advancement in Research Trends, SMART 2019* (pp. 266–270). <https://doi.org/10.1109/SMART46866.2019.9117512>

- Rahayu, A. N. S., Hermanto, T. I., & Nugroho, I. M. (2022). Sentiment analysis using K-Nearest Neighbor based on particle swarm optimization according to sunscreen's reviews. *Jurnal Teknik Informatika (JUTIF)*, 3(6), 1639–1646. <https://doi.org/10.20884/1.jutif.2022.3.6.425>
- Rukmana, D., Putri, A. N., Karim, A., & ZA, M. (2025). Sentiment classification of customer reviews in the fast-food industry using the Naïve Bayes algorithm. *bit-Tech*, 8(1), 930–940. <https://doi.org/10.32877/bt.v8i1.2766>
- Salsabillah, D. F., Ratnawati, D. E., & Setiawan, N. Y. (2024). Analisis sentimen ulasan rumah makan menggunakan perbandingan algoritma Support Vector Machine dengan Naive Bayes (Studi kasus: Ayam Goreng Nelongso Cabang Singosari, Malang). *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11(1), 107–116. <https://doi.org/10.25126/jtiik.20241117584>
- Samudra, B., Lestanti, S., & Romadhona, R. D. (2025). Sentiment analysis of Mie Gacoan reviews in Blitar City on the Grab application using the Support Vector Machine method. *JOSAR (Journal of Students Academic Research)*, 10(2), 129–139. <https://doi.org/10.35457/jwf5eb71>
- Sari, A. W., Hermanto, T. I., & Defriani, M. (2023). Sentiment analysis of tourist reviews using K-Nearest Neighbors algorithm and Support Vector Machine. *Sinkron: Jurnal dan Penelitian Teknik Informatika*, 7(3), 1366–1378. <https://doi.org/10.33395/sinkron.v8i3.12447>
- Sifa Amalia, B., Umaidah, Y., & Mayasari, R. (2021). Analisis sentimen review pelanggan restoran menggunakan algoritma Support Vector Machine dan K-Nearest Neighbor. *SITEKIN: Jurnal Sains, Teknologi dan Industri*, 19(1), 28–34.
- Singgale, Y. A. (2023). Analisis sentimen konsumen terhadap *food, services, and value* di restoran dan rumah makan populer Kota Makassar berdasarkan rekomendasi TripAdvisor menggunakan metode CRISP-DM dan SERVQUAL. *Building of Informatics, Technology and Science (BITS)*, 4(4). <https://doi.org/10.47065/bits.v4i4.3231>